



**HAL**  
open science

# Homogenous Multitype Fragmentations

Jean Bertoin

► **To cite this version:**

| Jean Bertoin. Homogenous Multitype Fragmentations. 2006. hal-00121940

**HAL Id: hal-00121940**

**<https://hal.science/hal-00121940>**

Preprint submitted on 22 Dec 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Homogenous Multitype Fragmentations

Jean Bertoin

*Laboratoire de Probabilités et Modèles Aléatoires  
Université Pierre et Marie Curie  
and DMA, Ecole Normale Supérieure  
Paris, France*

**Summary.** A homogeneous mass-fragmentation, as it has been defined in [6], describes the evolution of the collection of masses of fragments of an object which breaks down into pieces as time passes. Here, we show that this model can be enriched by considering also the types of the fragments, where a type may represent, for instance, a geometrical shape, and can take finitely many values. In this setting, the dynamics of a randomly tagged fragment play a crucial role in the analysis of the fragmentation. They are determined by a Markov additive process whose distribution depends explicitly on the characteristics of the fragmentation. As applications, we make explicit the connexion with multitype branching random walks, and obtain multitype analogs of the pathwise central limit theorem and large deviation estimates for the empirical distribution of fragments.

**Key words.** Multitype, fragmentation, branching process, Markov additive process.

**A.M.S. Classification.** 60J80, 60G18

**e-mail.** jbe@ccr.jussieu.fr

## 1 Introduction

In the recent years, there has been some interest for a class stochastic processes which are meant to serve as models for the evolution of an object that breaks down into smaller pieces, randomly and repeatedly as time passes. We refer to the monograph [6] and the survey [5] for a detailed account and references. Several crucial hypotheses have to be made in order to deal with models that can be analyzed by standard probabilistic techniques. Typically, one assumes that the process enjoys the branching property, in the sense that the dynamics of a given fragment do not depend on the others. A further important assumption which is made in [6], is that each fragment is characterized by a real number which can be viewed as its size. The latter requirement does not allow us to consider geometrical properties like the shape of a fragment, although such notions could be relevant for describing how an object breaks down.

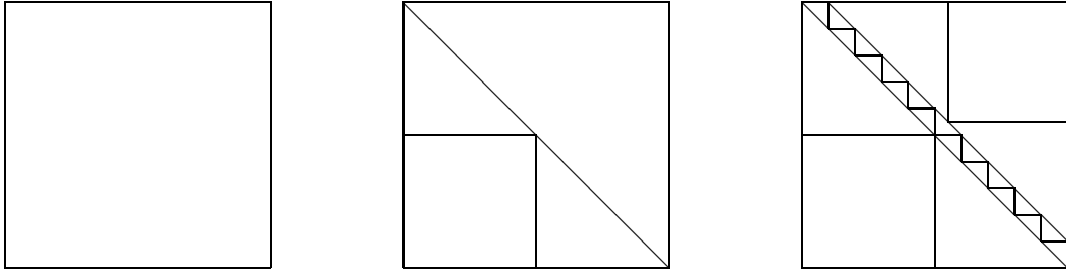


Figure 1: Example of a fragmentation of a square into squares and triangles

In the simpler case when time is discrete, one can analyze a fragmentation chain using the framework of branching random walks or that of multiplicative cascades. In this setting, it is therefore natural to enrich the model by assigning to each fragment a *type*, which, for instance, may describe its shape, and let the evolution of each fragment depend on its initial type. The study of the latter can then be developed directly by translating the literature on multitype branching random walks or cascades (see e.g. [2, 8, 9]).

However, we shall be interested here in the much more delicate case where time is continuous and each fragment may split immediately, a situation which cannot be handled directly by discrete techniques based on branching processes. In the monotype setting, Kingman's theory of exchangeable random partitions provides the key for the construction and the study of fragmentation processes in continuous time; this was pointed out first by Pitman [14], see Chapter 3 in [6] for a complete account. In the first part of this work (Sections 2 and 3), we shall briefly explain how Kingman's theory can be extended to the multitype setting (for any finite family of types), and how this extension enables us to develop an adequate framework for multitype fragmentation processes. In short, the main result states that the dynamics of a homogeneous multitype fragmentation are characterized by a family of erosion coefficients and a family of dislocation measures. Each erosion coefficient describes the rate at which a fragment with a given type melts down as time passes, and each dislocation measure specifies the statistics of its sudden splits. Once the correct setting is found, statements are straightforward modifications of that in the monotype situation, and for the sake of avoiding what would be essentially a lengthy and boring duplication of existing material, our presentation will be rather sketchy and proofs will be omitted. Non-specialist readers may wish to consult first Chapters 2 and 3 of [6] for getting the flavor of the arguments.

The second part of this work (Section 4) is devoted to the study of the tagged fragment, i.e. the fragment which contains a point which has been tagged at random and independently of the fragmentation process, and its applications. It departs more significantly from the monotype situation, in the sense that the evolution of the tagged fragment is now given in terms of a Markov additive process (instead of a subordinator), which depends explicitly on the characteristics of the fragmentation. The central limit theorem for Markov additive processes then enables us to determine the asymptotic behavior of certain multitype fragmentation processes, extending an old result of Kolmogorov [13] in this area. We will also develop the natural connexion with multitype branching random walks from which we derive some sharp large-deviation estimates based on the work of Biggins and Rahimzadeh Sani [9].

Throughout this text, we shall consider a finite family of types, say with cardinal  $k + 1 \geq 2$ , which can thus be identified with  $\{0, 1, \dots, k\}$ . The type 0 is special and will only be used in peculiar situations. As it has been mentioned above, it may be convenient to think of a type as a geometrical shape (see Figure 1 above for an example), but the type can also be used, for instance, to distinguish between active and inactive fragments in a frozen fragmentation (see [10] for a closely related notion in the setting of coalescents). Last but not least, it was observed recently by Haas *et al.* [11] that homogeneous fragmentations bear close connexions with certain continuum random trees, a class of random fractal spaces which has been introduced by Aldous. It is likely that more generally, multitype fragmentations can be used to construct some multifractal continuum random trees, following the analysis developed in [11].

## 2 Kingman's theory for partitions with types

The purpose of this section is to provide a brief presentation of an extension of Kingman's theory (see [12] or Section 2.3.2 in [6]) to partitions with types, which is a key step in the analysis of random fragmentations.

### 2.1 Partitions with types

We shall deal with two natural notions of partitions with types, which correspond to two different points of view. The first one focuses on the masses (and the types) of the components, whereas, roughly speaking, the second one corresponds to a discretization of the object which breaks down.

We call any numerical sequence  $\mathbf{x} = (x_1, x_2, \dots)$  with

$$x_1 \geq x_2 \geq \dots \geq 0 \quad \text{and} \quad \sum_1^\infty x_n \leq 1$$

a *mass-partition*, and write  $\mathcal{P}_m$  for the space of mass-partitions. A mass-partition with types is a pair  $\bar{\mathbf{x}} = (\mathbf{x}, \mathbf{i})$  with  $\mathbf{x} = (x_1, x_2, \dots) \in \mathcal{P}_m$  and  $\mathbf{i} = (i_1, i_2, \dots)$  a sequence in  $\{0, 1, \dots, k\}$ , such that for every  $n \in \mathbb{N} = \{1, 2, \dots\}$

$$x_n = 0 \Leftrightarrow i_n = 0, \tag{1}$$

and further

$$x_n = x_{n+1} \Rightarrow i_n \geq i_{n+1},$$

i.e. the sequence  $(x_1, i_1), (x_2, i_2), \dots$  is non-increasing in the lexicographic order. We shall write indifferently

$$\bar{\mathbf{x}} = (\mathbf{x}, \mathbf{i}) = ((x_1, i_1), (x_2, i_2), \dots)$$

by a slight abuse of notation.

We should think of  $x_n$  as the size of the  $n$ -th largest component of some object with total mass 1 which has been split, and of  $i_n$  as its type. A component with size 0 means that it is

absent or empty, and thus has the special type 0. Note that a mass-partition  $\mathbf{x}$  can be *improper*, in the sense that  $\sum_1^\infty x_n < 1$ . Then the mass-defect  $x_0 = 1 - \sum_1^\infty x_n$  is called the mass of *dust*, where the dust is viewed as a set of infinitesimal particles. It may be convenient to think that the special type 0 is also assigned to these infinitesimal particles.

We write  $\bar{\mathcal{P}}_m$  for the space of mass-partitions with types and endow it with the following distance. Let  $(e_1, \dots, e_k)$  denote the canonical basis of the Euclidean space  $\mathbb{R}^k$ , and associate to any mass-partition with types  $\bar{\mathbf{x}} \in \bar{\mathcal{P}}_m$  the probability measure on the axes of the unit cube

$$\varphi_{\bar{\mathbf{x}}} = x_0 \delta_0 + \sum_{n=1}^{\infty} x_n \delta_{x_n e_{i_n}},$$

where  $x_0 = 1 - \sum_1^\infty x_n$  is the mass of dust. Then we define the distance  $d(\bar{\mathbf{x}}, \bar{\mathbf{x}}')$  for every  $\bar{\mathbf{x}}, \bar{\mathbf{x}}' \in \bar{\mathcal{P}}_m$  as the Prohorov distance between the probability measures  $\varphi_{\bar{\mathbf{x}}}$  and  $\varphi_{\bar{\mathbf{x}}}'$ , which makes  $(\bar{\mathcal{P}}_m, d)$  a compact space. We stress that the distance  $d$  is strictly weaker than other perhaps simpler distances<sup>1</sup> on  $\bar{\mathcal{P}}_m$  such as, for instance,

$$\max\{|x_n - x'_n| \mathbb{1}_{\{i_n=i'_n\}} + (x_n + x'_n) \mathbb{1}_{\{i_n \neq i'_n\}} : n \in \mathbb{N}\}. \quad (2)$$

Next we turn our attention to the second notion of partition. We call any subset of  $\mathbb{N} = \{1, 2, \dots\}$  a block. A *partition* of a block  $B \subseteq \mathbb{N}$  is a sequence  $\pi = (\pi_1, \pi_2, \dots)$  of pairwise disjoint blocks with  $\cup \pi_n = B$ , which is ranked according to the increasing order of the least elements, i.e.  $\inf \pi_m \leq \inf \pi_n$  whenever  $m \leq n$  (with the usual convention that  $\inf \emptyset = \infty$ ). We write  $\mathcal{P}_B$  for the space of partitions of  $B$ .

Given a partition  $\pi = (\pi_1, \pi_2, \dots) \in \mathcal{P}_B$ , we can assign to each block  $\pi_n$  a type  $i_n \in \{0, \dots, k\}$ , with the following convention which is related to (1) :

$$i_n = 0 \Leftrightarrow \pi_n \text{ is either empty or a singleton.} \quad (3)$$

We write  $\bar{\pi} = (\pi, \mathbf{i}) = ((\pi_1, i_1), (\pi_2, i_2), \dots)$  and call  $\bar{\pi}$  a partition with types of  $B$ . We denote by  $\bar{\mathcal{P}}_B$  the space of partitions with types of some block  $B$ .

For every block  $B \subseteq \mathbb{N}$  and every partition  $\pi = (\pi_1, \dots)$  of  $\mathbb{N}$ , we define  $\pi|_B$ , the restriction of  $\pi$  to  $B$ , as the partition of  $B$  whose blocks are given by  $\pi_n \cap B$ ,  $n \in \mathbb{N}$ . If  $\bar{\pi} = (\pi, \mathbf{i})$  is now a partition with types, we assign types to the blocks of the restricted partition  $\pi|_B$  as follows. The type of  $\pi_n \cap B$  coincides with the type  $i_n$  of the block  $\pi_n$  if  $\pi_n \cap B$  is neither empty nor a singleton, and is 0 otherwise in order to agree with (3). We then write  $\bar{\pi}|_B$  for the restriction to  $B$  of the partition with types  $\bar{\pi}$ .

For every pair  $(\bar{\pi}, \bar{\pi}')$  of partitions with types, we define

$$d(\bar{\pi}, \bar{\pi}') = 1 / \sup\{n \in \mathbb{N} : \bar{\pi}|_{[n]} = \bar{\pi}'|_{[n]}\},$$

---

<sup>1</sup>It may be worthy to point out that our choice for the distance is well-adapted to the requirement (1). Typically, denote for  $n \in \mathbb{N}$  by  $\bar{\mathbf{x}}^{(n,i)}$  the mass-partition with types which consists in  $n$  identical fragments with mass  $1/n$  and fixed type  $i \in \{1, \dots, k\}$ . Then  $\bar{\mathbf{x}}^{(n,i)}$  converges as  $n \rightarrow \infty$  to the degenerate partition of pure dust (and type 0). Such a natural convergence would fail if we had chosen a stronger distance on  $\bar{\mathcal{P}}_m$  like (2).

where  $[n] = \{1, \dots, n\}$  and  $1/\sup \mathbb{N} = 0$ . Note that, as the type assigned to singletons is always 0, the identity  $\bar{\pi}|_{[1]} = \bar{\pi}'|_{[1]}$  holds in all cases and thus  $d(\bar{\pi}, \bar{\pi}') \leq 1$ . It is easily checked that  $d(\bar{\pi}, \bar{\pi}')$  defines a distance which makes  $\bar{\mathcal{P}}_{\mathbb{N}}$  a compact set; see Lemma 2.6 in [6] on its page 96.

Finally, we say that a block  $B \subseteq \mathbb{N}$  possesses an asymptotic frequency if and only if

$$|B| = \lim_{n \rightarrow \infty} n^{-1} \text{Card}(B \cap [n])$$

exists. If all the blocks of  $\bar{\pi} \in \bar{\mathcal{P}}_{\mathbb{N}}$  possess an asymptotic frequency, then we say that  $\bar{\pi}$  has asymptotic frequencies, and we write  $|\bar{\pi}|^{\downarrow} = (\mathbf{x}, \mathbf{i})$  for the sequence of the asymptotic frequencies and types of the blocks of  $\bar{\pi}$  ranked in the non-increasing lexicographic order. Note from Fatou's lemma that  $\sum_1^{\infty} |\pi_n| \leq 1$  is a mass-partition and thus  $|\bar{\pi}|^{\downarrow} \in \bar{\mathcal{P}}_{\mathbf{m}}$ .

## 2.2 Exchangeability and paintbox construction

A finite permutation is a bijection  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$  such that  $\sigma(n) = n$  when  $n$  is sufficiently large. The group of finite permutations acts naturally on the space  $\bar{\mathcal{P}}_{\mathbb{N}}$  of partitions with types. Specifically, we write  $\sigma^{-1}$  for the finite permutation obtained as the inverse  $\sigma$ . Given an arbitrary partition with types of  $\mathbb{N}$ ,  $\bar{\pi} = (\pi, \mathbf{i})$ ,  $\sigma^{-1}$  maps each block  $\pi_n$  of  $\pi$  into a block  $\sigma^{-1}(\pi_n)$  of some partition denoted by  $\sigma(\pi)$ . We decide to assign the type  $i_n$  of the block  $\pi_n$  to the block  $\sigma^{-1}(\pi_n)$ . This way, we obtain a partition with types denoted by  $\sigma(\bar{\pi})$ .

A measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$  is called *exchangeable* if it is invariant under the action of finite permutations. Following Kingman [12], we can associate to every mass-partition with types  $\bar{\mathbf{x}} = (\mathbf{x}, \mathbf{i}) \in \bar{\mathcal{P}}_{\mathbf{m}}$  an exchangeable probability measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$  by the paintbox construction. Specifically, introduce a pair of random variables  $(\xi, \tau)$  with values in  $\mathbb{Z}_+ \times \{0, 1, \dots, k\}$  whose distribution is specified by the following :

$$\mathbb{P}((\xi, \tau) = (n, i_n)) = x_n \text{ for every } n \in \mathbb{N} \text{ and } \mathbb{P}((\xi, \tau) = (0, 0)) = 1 - \sum_{n=1}^{\infty} x_n.$$

Then consider a sequence  $(\xi_1, \tau_1), \dots$  of i.i.d. copies of  $(\xi, \tau)$  and define a random partition with types  $\bar{\pi} = (\pi, \mathbf{i})$  by declaring that two distinct integers  $\ell, m$  are in the same block of  $\pi$  if and only if  $\xi_m = \xi_{\ell} \geq 1$ , and then decide that the type of that block is  $\tau_m = \tau_{\ell}$ . Integers  $\ell$  such that  $\xi_{\ell} = 0$  form the class of singletons of  $\pi$ , and their type is of course 0. Similarly, if some block of  $\pi$  is empty, then its type is necessarily 0 by our convention. The distribution of  $\bar{\pi}$  will be denoted by  $\rho_{\bar{\mathbf{x}}}$  and called the paintbox based on  $\bar{\mathbf{x}}$ .

A slight variation of this paintbox construction can be illustrated as follows. Suppose for simplicity that the mass-partition with types  $\bar{\mathbf{x}}$  can be represented by splitting some geometric object, for instance a rectangle with unit area, into smaller components, for instance squares, rectangles and triangles. Each component has an area and a shape which we called a type. Imagine that we pick at random a sequence of i.i.d. uniform points  $U_1, U_2, \dots$  in the initial object. A random partition with types is obtained by declaring that two distinct indices are in the same block of the partition if the corresponding random points belong to the same component of the object, and the type of this block is then the type of this component. See Figure 2 below.

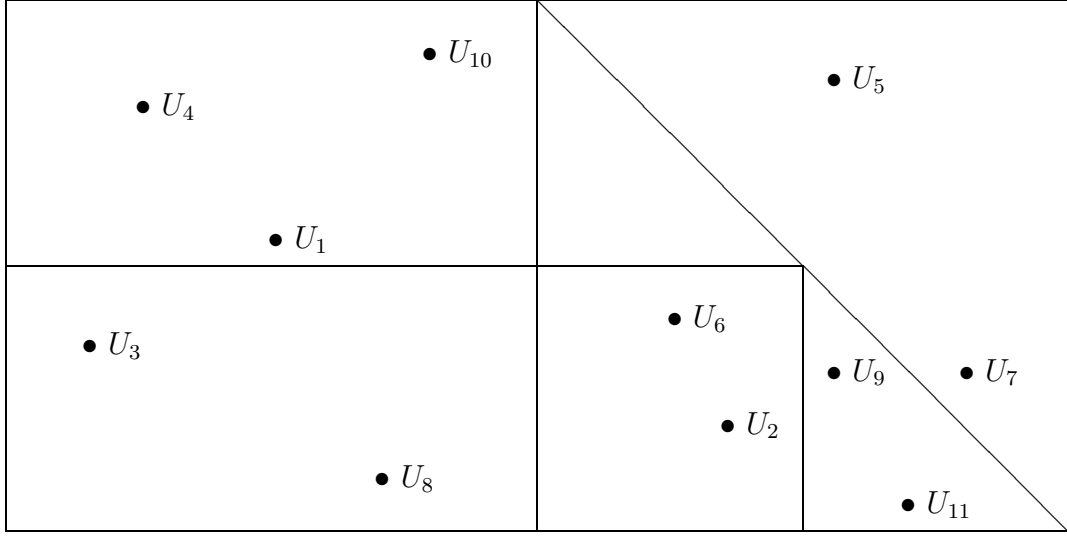


Figure 2 : paintbox with types for a partition of a rectangle;  
types : square = 1, rectangle = 2 and triangle = 3

$$\bar{\pi}_{[11]} = ((\{1, 4, 10\}, 2), (\{2, 6\}, 1), (\{3, 8\}, 2), (\{5, 7\}, 3), (\{9, 11\}, 3), (\emptyset, 0), \dots)$$

By the law of large numbers, for every positive integer  $\ell$ , the block  $B = \{m \in \mathbb{N} : \xi_m = \ell\}$  has an asymptotic frequency

$$|B| = \lim_{n \rightarrow \infty} n^{-1} \text{Card}(B \cap [n]) = x_\ell,$$

and clearly the type  $i_\ell$ . One can arrange the sequence of the pairs (asymptotic frequency, type) of the blocs of  $\bar{\pi}$  in the non-increasing lexicographic order and then it coincides with  $\bar{\mathbf{x}}$ .

Another important observation is that  $|\bar{\pi}_1|$ , the asymptotic frequency and the type of the first block  $\bar{\pi}_1$  of a paintbox based on a mass-partition with type  $\bar{\mathbf{x}}$ , has the distribution of a size-biased sample of  $\bar{\mathbf{x}}$ , that is

$$\rho_{\bar{\mathbf{x}}}(|\bar{\pi}_1| = (x_n, i_n)) = x_n \text{ for every } n \in \mathbb{N} \text{ and } \rho_{\bar{\mathbf{x}}}(|\bar{\pi}_1| = (0, 0)) = x_0 = 1 - \sum_1^\infty x_n. \quad (4)$$

Plainly, if  $\sigma$  is a finite permutation, then  $(\xi_{\sigma(1)}, \tau_{\sigma(1)}), (\xi_{\sigma(2)}, \tau_{\sigma(2)}), \dots$  is again a sequence of i.i.d. copies of  $(\xi, \tau)$  and the corresponding partition with types is given by  $\sigma(\bar{\pi})$ . Thus  $\rho_{\bar{\mathbf{x}}}$  is an exchangeable probability measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$ , and more generally any mixture of paintboxes produces an exchangeable probability measure on  $\bar{\mathcal{P}}_{\mathbf{m}}$ . The converse to the latter assertion is a slight variation of the fundamental theorem of Kingman [12], see e.g. Theorem 2.1 in [6] on its page 100.

**Theorem 1** *Let  $\rho$  be an exchangeable probability measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$ . Then  $\rho$ -almost every  $\bar{\pi} \in \bar{\mathcal{P}}_{\mathbb{N}}$  possesses asymptotic frequencies, and if  $\varrho$  stands for the distribution of  $|\bar{\pi}|^\downarrow$  under  $\rho$ , then there is the following disintegration of the measure  $\rho$  :*

$$\rho(d\bar{\pi}) = \int_{\bar{\mathcal{P}}_{\mathbb{m}}} \rho_{\bar{\mathbf{x}}}(d\bar{\pi}) \varrho(d\bar{\mathbf{x}}), \quad \bar{\pi} \in \bar{\mathcal{P}}_{\mathbb{N}}. \quad (5)$$

*Conversely, for every probability measure  $\varrho$  on  $\bar{\mathcal{P}}_{\mathbb{m}}$ , (5) defines an exchangeable probability measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$ .*

We next turn our attention to an extension of Kingman's theorem to certain sigma-finite measures on  $\bar{\mathcal{P}}_{\mathbb{N}}$ . In this direction, it is convenient to denote for every type  $i \in \{1, \dots, k\}$  and every block  $B$  that is neither empty nor a singleton, by  $\mathbf{1}_{B,i}$  the partition with type of  $B$  given by  $((B, i), (\emptyset, 0), \dots)$ . We also write  $\mathbf{1}_i = ((1, i), (0, 0), \dots) \in \bar{\mathcal{P}}_{\mathbb{m}}$  for a related mass-partition with types. For every  $n \in \mathbb{N}$ , we denote by  $\epsilon_{(n,i)}$  for the partition with types of  $\mathbb{N}$  which has exactly two non-empty blocks,  $(\mathbb{N} \setminus \{n\}, i)$  and  $(\{n\}, 0)$ . The exchangeable measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$

$$\epsilon_i = \sum_{n \in \mathbb{N}} \delta_{\epsilon_{(n,i)}}$$

will be referred to as the erosion measure with type  $i$ .

The following extension of Theorem 1 to certain possibly infinite measures is the multitype version of Theorem 3.1 in [6] on its page 127. Recall the notation  $\bar{\pi}|_B$  for the partition with types restricted to some block  $B$ , and that  $[2] = \{1, 2\}$ .

**Theorem 2** *Fix a type  $i \in \{1, \dots, k\}$  and let  $\mu_i$  be an exchangeable measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$  such that*

$$\mu_i(\{\mathbf{1}_{\mathbb{N},i}\}) = 0 \text{ and } \mu_i(\bar{\pi} \in \bar{\mathcal{P}}_{\mathbb{N}} : \bar{\pi}|_{[2]} \neq \mathbf{1}_{[2],i}) < \infty. \quad (6)$$

*Then the following holds:*

- (i)  $\mu_i$ -almost every partition  $\bar{\pi} \in \bar{\mathcal{P}}_{\mathbb{N}}$  possesses asymptotic frequencies.
- (ii) Let  $|\mu_i|^\downarrow$  be the image measure of  $\mu_i$  by the mapping  $\bar{\pi} \rightarrow |\bar{\pi}|^\downarrow$ . The restriction

$$\nu_i(d\bar{\mathbf{x}}) = \mathbb{1}_{\{\bar{\mathbf{x}} \neq \mathbf{1}_i\}} |\mu_i|^\downarrow(d\bar{\mathbf{x}}) \quad (7)$$

*of  $|\mu_i|^\downarrow$  to  $\bar{\mathcal{P}}_{\mathbb{m}} \setminus \{\mathbf{1}_i\}$  fulfills*

$$\int_{\bar{\mathcal{P}}_{\mathbb{m}}} (1 - x_1 \mathbb{1}_{\{i_1=i\}}) \nu_i(d\bar{\mathbf{x}}) < \infty \quad (8)$$

*and there is the disintegration*

$$\mathbb{1}_{\{|\bar{\pi}|^\downarrow \neq \mathbf{1}_i\}} \mu_i(d\bar{\pi}) = \int_{\bar{\mathcal{P}}_{\mathbb{m}}} \rho_{\bar{\mathbf{x}}}(d\bar{\pi}) \nu_i(d\bar{\mathbf{x}}).$$

- (iii) *There is a real number  $c_i \geq 0$  such that*

$$\mathbb{1}_{\{|\bar{\pi}|^\downarrow = \mathbf{1}_i\}} \mu_i(d\bar{\pi}) = c_i \epsilon_i(d\bar{\pi}). \quad (9)$$



Conversely, for every real number  $c_i \geq 0$  and every measure  $\nu_i$  on  $\bar{\mathcal{P}}_m$  without atom at  $\mathbf{1}_i$  and that satisfies (8), the measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$

$$\mu_i(d\bar{\pi}) = c_i \epsilon_i(d\bar{\pi}) + \int_{\bar{\mathcal{P}}_m} \rho_{\bar{\mathbf{x}}}(d\bar{\pi}) \nu_i(d\bar{\mathbf{x}})$$

is exchangeable and fulfills (6) and (7).

As the erosion measure  $\epsilon_i$  has infinite total mass, we see that  $c_i$  must be zero whenever  $\mu_i$  has a finite total mass. In this situation, Theorem 2 is an immediate consequence of Theorem 1.

### 3 The structure of multitype fragmentations

The purpose of this section is to describe the structure of multitype fragmentations. In the monotype case, dynamics of a homogeneous fragmentation are entirely determined by an erosion rate  $c \geq 0$ , which accounts for the smooth evolution of the process, and a dislocation measure  $\nu$  on the space  $\mathcal{P}_m$  of mass-partitions, which, as its name suggests, characterizes the statistics of the sudden dislocations. See Sections 3.1 and 3.2 in [6]. A similar description remains valid in the multitype situation, more precisely dynamics are then determined by a family  $(c_i)_{i \in \{1, \dots, k\}}$  of erosion rates and a family  $(\nu_i)_{i \in \{1, \dots, k\}}$  of dislocation measures on  $\bar{\mathcal{P}}_m$ , where the index  $i$  refers of course to the type of the fragment that is eroded or dislocated. This will be achieved first in the setting of partitions with types of  $\mathbb{N}$ , and then shifted to the more intuitive framework of mass-partitions.

#### 3.1 Basic definitions

We first introduce the natural notion of homogeneous fragmentation for mass-partitions with types, which bears strong similarities with that of multitype branching process. Specifically, let  $\bar{X} = (\bar{X}(t), t \geq 0)$  be a Markov process with values in  $\bar{\mathcal{P}}_m$  and càdlàg sample paths. For every  $i \in \{1, \dots, k\}$ , we write  $\mathbb{P}_i$  for its distribution starting from  $\bar{X}(0) = \mathbf{1}_i$ , i.e. at the initial time, there is a single unit mass with type  $i$ .

For every mass-partition with types  $\bar{\mathbf{x}} = (\mathbf{x}, \mathbf{i})$  and every real number  $r \geq 0$ , it will be convenient to write

$$r\bar{\mathbf{x}} = (r\mathbf{x}, \mathbf{i}) = ((rx_1, i_1), (rx_2, i_2), \dots).$$

We then introduce a sequence of independent processes  $Y^{(1)}, Y^{(2)}, \dots$  such that for every  $n \in \mathbb{N}$ ,  $Y^{(n)}$  is distributed as  $x_n \bar{X}$  under  $\mathbb{P}_{i_n}$ . For every  $t \geq 0$ , we write  $Y(t)$  for the rearrangement in the non-increasing lexicographic order of the terms of the random mass-partitions with types  $Y^{(1)}(t), Y^{(2)}(t), \dots$ , and denote by  $\mathbb{P}_{\bar{\mathbf{x}}}$  the distribution of the process  $Y = (Y(t), t \geq 0)$ . In particular  $\mathbb{P}_{\mathbf{1}_i} = \mathbb{P}_i$ .

**Definition 1** *The process  $\bar{X}$  is called a **homogeneous multitype mass-fragmentation** if, in the sense of the Markov property, the law of  $\bar{X}$  started from an arbitrary state  $\bar{\mathbf{x}} \in \bar{\mathcal{P}}_m$  is  $\mathbb{P}_{\bar{\mathbf{x}}}$ .*

The preceding section incites us to translate Definition 1 in the setting of partitions with types of  $\mathbb{N}$ . In this direction, the notion of fragmentation operator (see Definition 3.1 in [6] on its page 114) has a natural extension in the multitype setting.

Specifically, consider  $\pi \in \mathcal{P}_B$  a partition of some block  $B$  and  $\bar{\pi}^{(\cdot)} = (\bar{\pi}^{(n)}, n \in \mathbb{N})$  a sequence of partitions with types. We then write  $\text{Frag}(\pi, \bar{\pi}^{(\cdot)})$  for the partition with types which is obtained from the collection of blocks with types of the sequence of the restrictions  $\bar{\pi}_{|\pi_n}^{(n)}$  of  $\bar{\pi}^{(n)}$  to the  $n$ -th block  $\pi_n$  of  $\pi$  for  $n \in \mathbb{N}$ , by rearrangement in the non-increasing lexicographic order. In other words, each block  $\pi_n$  of  $\pi$  is split using  $\bar{\pi}^{(n)}$ . Note that if the block  $\pi_n$  is either a singleton or empty, then the partition with types  $\bar{\pi}_{|\pi_n}^{(n)}$  does not depend on  $\bar{\pi}^{(n)}$ ; more precisely, it is always given by  $((\pi_n, 0), (\emptyset, 0), (\emptyset, 0), \dots)$ .

Let  $\bar{\Pi} = (\bar{\Pi}(t), t \geq 0)$  be a Markov process with values in  $\bar{\mathcal{P}}_{\mathbb{N}}$  with càdlàg sample paths. By a slight abuse of notation, for every  $i \in \{1, \dots, k\}$ , we write  $\mathbb{P}_i$  for its distribution starting from  $\bar{\Pi}(0) = \mathbf{1}_{\mathbb{N}, i}$ .

**Definition 2** *The process  $\bar{\Pi}$  is called a **homogeneous multitype fragmentation** if for every time  $t \geq 0$  and every type  $i \in \{1, \dots, k\}$  the distribution of  $\bar{\Pi}(t)$  under  $\mathbb{P}_i$  is exchangeable, and the semigroup of  $\bar{\Pi}$  can be described as follows :*

*Fix  $t, t' \geq 0$  and consider a partition with types  $\bar{\pi} = (\pi, \mathbf{i})$  where  $\mathbf{i} = (i_1, i_2, \dots)$  is a sequence in  $\{0, 1, \dots, k\}$ . Let  $\bar{\pi}^{(\cdot)} = (\pi^{(1)}, \dots)$  denote a sequence of independent exchangeable random partitions with types, such that for every  $n \in \mathbb{N}$  with  $i_n \neq 0$ ,  $\bar{\pi}^{(n)}$  is distributed as  $\bar{\Pi}(t')$  under  $\mathbb{P}_{i_n}$ . When  $i_n = 0$ , the block  $\pi_n$  is either empty or a singleton; the role of  $\bar{\pi}^{(n)}$  has no importance and its law can be chosen arbitrarily. The conditional distribution of  $\bar{\Pi}(t+t')$  given  $\bar{\Pi}(t) = (\pi, \mathbf{i})$  is then the law of  $\text{Frag}(\pi, \bar{\pi}^{(\cdot)})$ .*

Let us now explain the connexion between these two definitions. When  $\bar{\Pi}$  is a homogeneous multitype fragmentation, we know from Kingman's Theorem 1 that for every  $t \geq 0$ , the exchangeable random partition with types  $\bar{\Pi}(t)$  possesses asymptotic frequencies a.s. If we write  $\bar{X}(t) = |\bar{\Pi}(t)|^\downarrow$  for the random multitype mass-partition obtained by reordering these asymptotic frequencies in the non-increasing lexicographic order, it can be proved that the process  $\bar{X} = (\bar{X}(t), t \geq 0)$  is then a homogeneous multitype mass-fragmentation. Technically, the main difficulty is to establish that the paths of  $t \rightarrow |\bar{\Pi}(t)|^\downarrow$  are càdlàg; in this direction we stress that this could fail if we had equipped  $\bar{\mathcal{P}}_{\mathbb{m}}$  with a stronger distance such as that given by (2). In the converse direction, one can rephrase the argument of Berestycki [3] and show that given a homogeneous multitype mass-fragmentation  $\bar{X} = (\bar{X}(t), t \geq 0)$ , there exists a homogeneous multitype fragmentation  $\bar{\Pi}$  such that the process  $(|\bar{\Pi}(t)|^\downarrow, t \geq 0)$  is distributed as  $\bar{X}$ . In short, there a bijective correspondence between the laws of homogeneous multitype mass-fragmentations and laws of homogeneous multitype fragmentations.

Of course, the fundamental difference with the monotype case (see for instance Definition 3.2 in [6] on its page 119) is that the distribution of the sequence  $\bar{\pi}^{(\cdot)}$  which is used to split the partition with types  $\bar{\Pi}(t)$  into finer the blocks depends on  $\bar{\Pi}(t)$ . However, this dependence only arises through the types of the blocks of  $\bar{\Pi}(t)$ , and does not involve directly the partition  $\bar{\Pi}(t)$ . This preserves the possibility of adapting the approach developed in Chapter 3 of [6], provided that one can handle some technical issues.

In particular, it is easily seen that the fragmentation operation is compatible with the restriction of partitions with types, in the sense that for every integer  $n$

$$\text{Frag}(\pi, \bar{\pi}^{(\cdot)})|_{[n]} = \text{Frag}(\pi|_{[n]}, \bar{\pi}^{(\cdot)}) . \quad (10)$$

This entails that the Markov property still holds for the restricted process  $\bar{\Pi}|_{[n]} = (\bar{\Pi}|_{[n]}(t), t \geq 0)$ , and as the latter can only take finitely many values,  $\bar{\Pi}|_{[n]}$  is a Markov chain in continuous times. Note that  $\bar{\Pi}|_{[n]}$  coincides with the restriction of  $\bar{\Pi}|_{[n+1]}$  to  $[n]$ , and that the initial process  $\bar{\Pi}$  can be recovered from the sequence of Markov chains  $\bar{\Pi}|_{[n]}$ ,  $n \in \mathbb{N}$ .

Just as in Section 3.1.2 of [6], these observations enable us to characterize the law of  $\bar{\Pi}$  by a finite family of measures  $(\mu_i)_{i \in \{1, \dots, k\}}$  on  $\bar{\mathcal{P}}_{\mathbb{N}}$  as follows. For every fixed  $i \in \{1, \dots, k\}$ ,  $n \geq 2$  and every partition with types  $\bar{\gamma}$  of  $[n]$  with  $\bar{\gamma} \neq \mathbf{1}_{[n], i}$ , we introduce the jump rate of  $\bar{\Pi}|_{[n]}$  from  $\mathbf{1}_{[n], i}$  to  $\bar{\gamma}$ ,

$$q_{n, \bar{\gamma}} = \lim_{t \rightarrow 0^+} \frac{1}{t} \mathbb{P}_i (\bar{\Pi}|_{[n]}(t) = \bar{\gamma}) .$$

By the very same arguments as in Section 3.1.2 of [6], one can check that the collection of those jump rates characterize the evolution of the restricted Markov chains  $\bar{\Pi}|_{[n]}$ , and thus of the process  $\bar{\Pi}$ . Further these jump rates can be represented as

$$q_{n, \bar{\gamma}} = \mu_i (\{ \bar{\pi} \in \bar{\mathcal{P}}_{\mathbb{N}} : \bar{\pi}|_{[n]} = \bar{\gamma} \}) , \quad (11)$$

where  $\mu_i$  is an exchangeable measure on  $\bar{\mathcal{P}}_{\mathbb{N}}$  with

$$\mu_i(\{\mathbf{1}_{\mathbb{N}, i}\}) = 0 \text{ and } \mu_i(\{ \bar{\pi} \in \bar{\mathcal{P}}_{\mathbb{N}} : \bar{\pi}|_{[n]} \neq \mathbf{1}_{[n], i} \}) < \infty \text{ for any } n \geq 2 ,$$

and these requirements determine the measure  $\mu_i$  uniquely. Note that when the condition above is fulfilled for  $n = 2$ , then, thanks to the exchangeability, it is fulfilled for every  $n \geq 2$ , therefore it is equivalent to (6). We shall refer to the family  $(\mu_i)_{i \in \{1, \dots, k\}}$  as the splitting rates of  $\bar{\Pi}$ .

## 3.2 Poissonian constructions

Our first goal in this section is to show that any family of exchangeable measures  $(\mu_i)_{i \in \{1, \dots, k\}}$  which fulfill (6) can be viewed as the splitting rates of a homogeneous multitype fragmentation  $\bar{\Pi}$ . More precisely, we shall briefly present a Poissonian construction of  $\bar{\Pi}$  which mimics that in Section 3.1.3 of [6] in the monotype case. For the sake of simplicity, we assume that the initial state has been chosen equal to  $\mathbf{1}_{\mathbb{N}, i_0}$  for some type  $i_0 \in \{1, \dots, k\}$ .

For every type  $i \in \{1, \dots, k\}$ , consider the atoms  $(t_{i, m}, \bar{\pi}^{(i, m)}, \ell_{i, m})_{m \in \mathbb{N}}$  of a Poisson random measure in  $\mathbb{R}_+ \times \bar{\mathcal{P}}_{\mathbb{N}} \times \mathbb{N}$  with intensity  $dt \otimes \mu_i \otimes \#$ , where  $\#$  stands for the counting measure on  $\mathbb{N}$ . This means that for every measurable set  $A \subseteq \mathbb{R}_+ \times \bar{\mathcal{P}}_{\mathbb{N}} \times \mathbb{N}$ , the cardinal of the collection of indices  $m$  for which  $(t_{i, m}, \bar{\pi}^{(i, m)}, \ell_{i, m}) \in A$  has the Poisson distribution with parameter  $dt \otimes \mu_i \otimes \#(A)$  and to disjoint sets correspond independent Poisson variables. We assume that these Poisson random measures are independent for different types  $i \in \{1, \dots, k\}$ .

For every integer  $n$ , we can construct a Markov chain in continuous time  $\bar{\Pi}^{[n]} = (\bar{\Pi}^{[n]}(t), t \geq 0)$  with values in  $\bar{\mathcal{P}}_{[n]}$  as follows. The atoms  $(t, \bar{\pi}, \ell)$  of the Poisson point measures such that

$\bar{\pi}_{[n]} = \mathbf{1}_{[n],i}$  or  $\ell > n$  play no role in the construction and can thus be removed. Thanks to (6), the instants  $t$  at which an atom  $(t, \bar{\pi}, \ell)$  that has not been removed arises, form a discrete set of  $\mathbb{R}_+$ , and the chain  $\bar{\Pi}^{[n]}$  can only jump at such times. More precisely, if  $(t_{i,m}, \bar{\pi}^{(i,m)}, \ell_{i,m})$  is an atom which has not been removed, then we look at the  $\ell_{i,m}$ -th block of  $\bar{\Pi}^{[n]}(t_{i,m}-)$ , say  $B$ . If the type of this block is different from the type  $i$  of the atom (in particular if  $B$  is either empty or a singleton), then we decide that  $\bar{\Pi}^{[n]}(t_{i,m}) = \bar{\Pi}^{[n]}(t_{i,m}-)$  and  $t_{i,m}$  is not a jump time for the chain  $\bar{\Pi}^{[n]}$ . If the type of  $B$  is the same as the type  $i$  of the atom, then  $\bar{\Pi}^{[n]}(t_{i,m})$  is the partition obtained from  $\bar{\Pi}^{[n]}(t_{i,m}-)$  by replacing  $B$ , that is the  $\ell_{i,m}$ -th block of  $\bar{\Pi}^{[n]}(t_{i,m}-)$ , by the restriction of  $\bar{\pi}^{(i,m)}$  to this block, and leaving the other blocks and types unchanged.

To give an example, take for instance  $n = 7$ ,  $\ell_{i,m} = 2$ ,

$$\bar{\pi}^{(i,m)} = ((\{1, 3, 5, 7, \dots\}, 3), (\{2, 4, 6, \dots\}, 1), \dots) ,$$

and set for simplicity  $t = t_{i,m}$ . Assume also that

$$\bar{\Pi}^{[n]}(t-) = ((\{1, 2\}, 1), (\{3, 4, 5\}, i), (\{6, 7\}, 2), (\emptyset, 0), \dots) .$$

As  $\ell_{i,m} = 2$ , we look at the 2nd block of  $\bar{\Pi}^{[n]}(t-)$ , which is  $B = \{3, 4, 5\}$  and has type  $i$ , and thus coincides with the type the atom  $(t_{i,m}, \bar{\pi}^{(i,m)}, \ell_{i,m})$ . At time  $t$ , we split  $B$  using the partition with types  $\bar{\pi}^{(i,m)}$ . This produces two new blocks with types:  $\{3, 5\}$  which has type 3, and  $\{4\}$  which is a singleton and thus has type 0. We conclude that

$$\bar{\Pi}^{[n]}(t) = ((\{1, 2\}, 1), (\{3, 5\}, 3), (\{4\}, 0), (\{6, 7\}, 2), (\emptyset, 0), \dots) .$$

It is easily seen that this construction is compatible with the restriction, in the sense that for every  $n \in \mathbb{N}$ ,  $\bar{\Pi}^{[n]}$  coincides with the restriction of  $\bar{\Pi}^{[n+1]}$  to  $[n]$ . We refer to Lemma 3.3 in [6] on its page 118 for the argument in the monotype case. This implies that there exists a process  $\bar{\Pi}$  with values in  $\bar{\mathcal{P}}_{\mathbb{N}}$  such that the restriction of  $\bar{\Pi}$  to  $[n]$  is  $\bar{\Pi}^{[n]}$ ; see Lemma 2.5 in [6] on its page 95 for a closely related argument.

A crucial step is to show that for every  $t \geq 0$ , the distribution of  $\bar{\Pi}(t)$  is exchangeable. The proof relies on the following technical lemma, which is the multitype version of Lemma 3.2 in [6] on its page 116.

**Lemma 1** *Let  $\bar{\pi} = (\pi, \mathbf{i}) \in \bar{\mathcal{P}}_{\mathbb{N}}$  be an exchangeable random partition with types and  $\bar{\pi}^{(\cdot)} = (\bar{\pi}^{(n)}, n \in \mathbb{N})$  a sequence of random partitions with types. Suppose that :*

- $\pi$  and  $\bar{\pi}^{(\cdot)}$  are independent conditionally on  $\mathbf{i}$ ,
- the sequence  $\bar{\pi}^{(\cdot)}$  is **doubly-exchangeable**, in the sense that for every finite permutation  $\sigma$  of  $\mathbb{N}$ , the sequences

$$(\sigma(\bar{\pi}^{(n)}), n \in \mathbb{N}) \quad \text{and} \quad (\bar{\pi}^{(\sigma(n))}, n \in \mathbb{N})$$

*both have the same law as  $\bar{\pi}^{(\cdot)}$ . Then the random partitions with types  $\bar{\pi}$  and  $\text{Frag}(\pi, \bar{\pi}^{(\cdot)})$  are jointly exchangeable, that is their joint distribution is invariant by the action of permutations.*

**Sketch of the proof:** One observes that with probability one, the conditional distribution of  $\pi$  given  $\mathbf{i}$  is an exchangeable probability measure on  $\mathcal{P}_{\mathbb{N}}$ . One can then follow the argument of the proof of Lemma 3.2 in [6].  $\square$

It is then easy to verify from standard properties of Poisson random measures that the process  $\bar{\Pi}^{[n]}$  which has just been constructed is a Markov chain in continuous time, and that its jumps rates

$$q_{n,\bar{\gamma}} = \lim_{t \rightarrow 0^+} \frac{1}{t} \mathbb{P}_i (\bar{\Pi}^{[n]}(t) = \bar{\gamma})$$

for every  $\bar{\gamma} \in \bar{\mathcal{P}}_{[n]}$  with  $\bar{\gamma} \neq \mathbf{1}_{[n],i}$ , are given by (11). This shows that the process  $\bar{\Pi}$ , which is specified by the requirement that its restriction to  $[n]$  coincides with  $\bar{\Pi}^{[n]}$ , is a homogeneous multitype fragmentation with splitting rates  $(\mu_i)_{i \in \{1, \dots, k\}}$ . Applying Theorem 2, we can summarize this analysis in the following statement.

**Proposition 1** *Let  $\bar{\Pi}$  be a homogeneous multitype fragmentation. There exists a unique family  $(c_i)_{i \in \{1, \dots, k\}}$  of nonnegative real numbers and a unique family  $(\nu_i)_{i \in \{1, \dots, k\}}$  of measures on  $\bar{\mathcal{P}}_m$  which fulfill (8), such that the family  $(\mu_i)_{i \in \{1, \dots, k\}}$  of splitting rates of  $\bar{\Pi}$  is given by (9).*

*Conversely, for every family  $(c_i)_{i \in \{1, \dots, k\}}$  of nonnegative real numbers and every family  $(\nu_i)_{i \in \{1, \dots, k\}}$  of measures on  $\bar{\mathcal{P}}_m$  which fulfill (8), if we define measures  $\mu_i$  on  $\bar{\mathcal{P}}_{\mathbb{N}}$  by (9), then the Poissonian construction above produces a homogeneous multitype fragmentation  $\bar{\Pi}$  with splitting rates  $(\mu_i)_{i \in \{1, \dots, k\}}$ .*

Berestycki [3] established a related Poissonian construction for monotype mass-fragmentations. The latter can be extended to the multitype setting provided that the erosion coefficients  $c_i$  are all the same, which enlighten the probabilistic interpretation of the dislocation measures  $\nu_i$ .

Specifically, for each type  $i \in \{1, \dots, k\}$ , consider the atoms  $(t_{i,m}, \bar{\mathbf{x}}^{(i,m)}, \ell_{i,m})_{m \in \mathbb{N}}$  of a Poisson random measure in  $\mathbb{R}_+ \times \bar{\mathcal{P}}_m \times \mathbb{N}$  with intensity  $dt \otimes \nu_i \otimes \#$ , where  $\#$  stands for the counting measure on  $\mathbb{N}$ . Assume as usual that these Poisson measures are independent for different types. One can construct a pure jump process  $(\bar{Y}(t), t \geq 0)$  in  $\bar{\mathcal{P}}_m$  which jumps only at times  $t_{i,m}$  at which some atom  $(t_{i,m}, \bar{\mathbf{x}}^{(i,m)}, \ell_{i,m})$  occurs. The jump (i.e. the dislocation) induced by such an atom can be described as follows.

We consider the mass-partition with types immediately before time  $t_{i,m}$ , that is  $\bar{Y}(t_{i,m}-)$ , and look at its  $\ell_{i,m}$ -th term, say  $(y, j)$  for some  $y \geq 0$  and  $j \in \{0, \dots, k\}$  (recall that the terms of a mass-partition with types are ranked in the non-increasing lexicographic order). If the type  $j$  is different from the type  $i$  of the atom, then we simply set  $\bar{Y}(t_{i,m}-) = \bar{Y}(t_{i,m})$ . Otherwise, the  $\ell_{i,m}$ -th term of  $\bar{Y}(t_{i,m}-)$  is dislocated according to  $\bar{\mathbf{x}}^{(i,m)}$ , that is it is replaced by the mass-partition with types  $y\bar{\mathbf{x}}^{(i,m)}$ . The other terms of  $\bar{Y}(t_{i,m}-)$  are left unchanged, and  $\bar{Y}(t_{i,m})$  then results from the rearrangement in the non-increasing lexicographic order of all the terms.

For instance, if

$$\ell_{i,m} = 2, \quad \bar{\mathbf{x}}^{(i,m)} = \left( \left( \frac{2}{3}, 2 \right), \left( \frac{1}{3}, 1 \right), (0, 0), \dots \right)$$

and

$$\bar{Y}(t_{i,m}-) = \left( \left( \frac{1}{2}, 4 \right), \left( \frac{1}{3}, i \right), \left( \frac{1}{6}, 1 \right), (0, 0), \dots \right)$$

then at time  $t_{i,m}$  the second term of  $\bar{Y}(t_{i,m}-)$ , i.e.  $(\frac{1}{3}, i)$  is dislocated using  $\bar{\mathbf{x}}^{(i,m)}$ . This produces the sequence  $((\frac{2}{9}, 2), (\frac{1}{9}, 1), (0, 0), \dots)$ , and finally

$$\bar{Y}(t_{i,m}) = \left( \left( \frac{1}{2}, 4 \right), \left( \frac{2}{9}, 2 \right), \left( \frac{1}{6}, 1 \right), \left( \frac{1}{9}, 1 \right), (0, 0), \dots \right).$$

The process  $\bar{Y}$  is then a homogeneous multitype fragmentation with zero erosion and dislocation measures  $(\nu_i)_{i \in \{1, \dots, k\}}$ . Following an argument in Berestycki [3], one can check that for every  $\mathbf{c} \geq 0$ , the exponentially discounted process  $\bar{X}(t) = e^{-\mathbf{c}u} \bar{Y}(t)$ ,  $t \geq 0$ , is then a homogeneous multitype mass-fragmentation with dislocation measures  $(\nu_i)_{i \in \{1, \dots, k\}}$  and erosion coefficients  $c_i = \mathbf{c}$  for every  $i \in \{1, \dots, k\}$ . Unfortunately, this simple transformation cannot be extended to the case when the erosion coefficients are distinct. Informally, when the erosion coefficients depend on the type of the fragments, one would need information about the types of the ancestors of each fragment of  $\bar{Y}(t)$  in order to determine the proportion of its mass that has been turned to dust at time  $t$ . This information is available for processes with values in  $\bar{\mathcal{P}}_{\mathbb{N}}$ , but not for those with values in  $\bar{\mathcal{P}}_{\mathbf{m}}$ .

## 4 The tagged fragment

Up to a few technical issues, the analysis of multitype fragmentations was so far an easy translation of that in the monotype situation. However more significant differences appear when dealing with finer aspects of these processes. Here, we shall focus on the evolution of the tagged fragment, that is the fragment which contains a point which has been picked uniformly at random and independently of the fragmentation process. The relevance of this study stems from the fact that, even though the tagged fragment alone does not characterize the evolution of the fragmentation, it captures some useful information. In particular, this will enable us to determine the asymptotic behavior of the fragmentation, by making explicit the connexion with multitype branching random walks.

Let  $\bar{\mathbf{X}} = (X, T)$  be a homogeneous multitype mass-fragmentation, where  $X_n(t)$  stands for the mass of the  $n$ -th largest fragment at time  $t$  and  $T_n(t)$  for its type. It will be convenient to think of  $\bar{\mathbf{X}}$  as associated to a homogeneous multitype fragmentation  $\bar{\Pi}$  by  $\bar{\mathbf{X}} = |\bar{\Pi}|^\downarrow$ . In order to avoid technical discussions, we shall assume throughout this section that the fragmentation process is *conservative*, i.e. for every type  $i \in \{1, \dots, k\}$ , the erosion coefficient is  $c_i = 0$  and the dislocation measure satisfies

$$\nu_i \left( \left\{ \bar{\mathbf{x}} \in \bar{\mathcal{P}}_{\mathbf{m}} : x_0 = 1 - \sum_1^\infty x_n > 0 \right\} \right) = 0. \quad (12)$$

The description of the evolution of the tagged fragment relies on the notion of *Markov additive processes*. We first provide some background in this area, referring to Section XI.2 of Asmussen [1] for details.

## 4.1 Background on Markov additive processes

The class of Markov additive processes that will be useful in this work is that formed by bivariate Markov processes  $(J_t, S_t)_{t \geq 0}$ , where  $(J_t)_{t \geq 0}$  is a continuous time Markov chain with values in the finite space of types  $\{1, \dots, k\}$ , and, roughly speaking, on every time-interval on which  $J$  stays constant,  $S$  evolves as a subordinator (i.e. an increasing process with independent and stationary increments) with characteristics specified by the value of  $J$ . More precisely, one requires that for every  $t, t' \geq 0$ ,

$$\mathbb{E}((f(S_{t+t'}) - f(S_t))g(J_{t+t'}) \mid J_t, S_t) = \mathbb{E}_{J_t, 0}(f(S_{t'})g(J_{t'})) \quad (13)$$

where  $\mathbb{E}_{j,s}$  refers to the mathematical expectation when the process  $(J, S)$  starts from the state  $(j, s)$  and  $f, g$  denote two generic measurable nonnegative functions.

The law of the Markov chain  $(J_t)_{t \geq 0}$  is specified by its intensity matrix  $\mathbf{\Lambda} = (\lambda_{ij})_{i,j \in \{1, \dots, k\}}$ , i.e. for  $i \neq j$ ,  $\lambda_{ij}$  is the jump rate of  $J$  from  $i$  to  $j$  and  $\sum_{j=1}^k \lambda_{ij} = 0$ . On every time-interval  $[t, t + t'[,$  on which  $J \equiv i$ ,  $S$  evolves as a subordinator  $S^{(i)}$  with Bernstein <sup>2</sup> exponent  $\Psi^{(i)}$ , i.e.

$$\mathbb{E}(\exp(-\theta S_t^{(i)})) = \exp(-t\Psi^{(i)}(\theta)).$$

The Bernstein exponent is a concave increasing function which can take the value  $-\infty$ , and is nonnegative and finite on  $[0, \infty[$ .

Further, a jump of  $J$  from  $i$  to  $j \neq i$  has a probability  $p_{ij}$  of inducing a jump of  $S$  at the same time, the distribution of which is denoted by  $B_{ij}$ , and we write  $\hat{B}_{ij}(\theta) = \int e^{-\theta x} B_{ij}(dx)$ . It is also convenient to agree that  $p_{ii} = 0$ .

For every types  $i, j \in \{1, \dots, k\}$  and every  $\theta \geq 0$  and  $t \geq 0$ , there is the following identity between  $k \times k$  matrices :

$$\mathbb{E}_{i,0}(e^{-\theta S_t}, J_t = j) = (e^{-t\Phi(\theta)})_{ij},$$

where

$$\Phi(\theta) = -\mathbf{\Lambda} + (\Psi^{(i)}(\theta))_{\text{diag}} + (\lambda_{ij} p_{ij} (1 - \hat{B}_{ij}(\theta))), \quad (14)$$

see Proposition 2.2 in [1] on its page 311. We shall refer to  $\Phi$  as the *Bernstein matrix* of  $(J, S)$ .

## 4.2 Distribution of the tagged fragment

We are interested in the process of the asymptotic frequency and the type of the first block  $(|\bar{\Pi}_1(t)|, t \geq 0)$  of a homogeneous multitype fragmentation  $\bar{\Pi}$ . Recall from the paintbox construction that  $|\bar{\Pi}_1(t)|$  can be viewed as the mass and type of the fragment which contains some point that has been picked at random according to the mass-distribution and independently of the fragmentation.

The conditions which have been enforced at the beginning of this section ensure that for every  $t \geq 0$ , the first block  $\Pi_1(t)$  is neither empty nor a singleton, hence its asymptotic frequency is

---

<sup>2</sup>Note that, since we are dealing with subordinators, we shall work with the Bernstein exponent  $\Psi^{(i)}$  whereas Asmussen [1] uses the cumulant  $\theta \rightarrow -\Psi^{(i)}(-\theta)$ .

strictly positive and its type is not 0. This allows us to introduce the process  $(J, S)$  with values in  $\{1, \dots, k\} \times \mathbb{R}_+$  by

$$|\bar{\Pi}_1(t)| = (\exp(-S_t), J_t), \quad t \geq 0.$$

**Theorem 3** *Suppose that the homogeneous multitype fragmentation  $\bar{\Pi}$  has erosion coefficients  $c_i = 0$  and that its dislocation measures fulfill (12). Then  $(J, S)$  is a Markov additive process with Bernstein matrix given for every  $\theta \geq 0$  by*

$$\Phi(\theta) = \left( \int_{\bar{\mathcal{P}}_m} \left( \mathbb{1}_{\{i=j\}} - \sum_{n=1}^{\infty} x_n^{1+\theta} \mathbb{1}_{\{i_n=j\}} \right) \nu_i(d\bar{\mathbf{x}}) \right)_{i,j \in \{1, \dots, k\}}.$$

**Proof:** The fact that  $(J, S)$  is a Markov process that satisfies (13) can be seen from the Poissonian construction and the arguments in Section 3.2.2 of [6].

The determination of the Bernstein matrix also relies on the Poissonian construction. First, note that for  $i \neq j$ , the jump rate  $\lambda_{ij}$  of the type process  $J$  coincides with the rate of occurrence of atoms  $(t_{i,m}, \bar{\pi}^{(i,m)}, 1)$  with  $\bar{\pi}^{(i,m)} = (\pi^{(i,m)}, \mathbf{i})$  and  $i_1 = j$ . Using (4) and Theorem 2, this yields

$$\lambda_{ij} = \int_{\bar{\mathcal{P}}_m} \left( \sum_{n=1}^{\infty} x_n \mathbb{1}_{\{i_n=j\}} \right) \nu_i(d\bar{\mathbf{x}}), \quad \text{for } i \neq j.$$

As  $\sum_{j=1}^k \lambda_{ij} = 0$ , this entails

$$\lambda_{ii} = - \int_{\bar{\mathcal{P}}_m} \left( \sum_{n=1}^{\infty} x_n \mathbb{1}_{\{i_n \neq i\}} \right) \nu_i(d\bar{\mathbf{x}}).$$

Thus, by (12), we obtain the general formula

$$\lambda_{ij} = \int_{\bar{\mathcal{P}}_m} \left( \sum_{n=1}^{\infty} x_n \mathbb{1}_{\{i_n=j\}} - \mathbb{1}_{\{i=j\}} \right) \nu_i(d\bar{\mathbf{x}}), \quad i, j \in \{1, \dots, k\}. \quad (15)$$

A slight refinement of this argument enables us to compute the finite measure  $\lambda_{ij} p_{ij} B_{ij}$ . Specifically, one finds for  $i \neq j$

$$\lambda_{ij} p_{ij} \int f(b) B_{ij}(db) = \int_{\bar{\mathcal{P}}_m} \left( \sum_{n=1}^{\infty} x_n \mathbb{1}_{\{i_n=j\}} f(-\ln x_n) \right) \nu_i(d\bar{\mathbf{x}}).$$

This gives

$$\lambda_{ij} p_{ij} (1 - \hat{B}_{ij}(\theta)) = \int_{\bar{\mathcal{P}}_m} \left( \sum_{n=1}^{\infty} \mathbb{1}_{\{i_n=j\}} (x_n - x_n^{1+\theta}) \right) \nu_i(d\bar{\mathbf{x}}).$$

Finally, the calculation of the Bernstein functions of the subordinators  $S^{(i)}$  is made by reduction to the monotype situation. Specifically, we shall work under the law  $\mathbb{P}_i$ , and we denote by  $\nu_i^\dagger$  the image of  $\nu_i$  by the map  $\dagger_i : \bar{\mathcal{P}}_m \rightarrow \mathcal{P}_m$  where  $\dagger_i(\bar{\mathbf{x}})$  is the mass-partition given



by rearrangement of the terms  $\mathbb{1}_{\{i_n=i\}}x_n$ . Informally, this means that all the components of  $\bar{\mathbf{x}}$  which are not of type  $i$  are reduced to dust. Then  $\nu_i^\dagger$  is a (monotype) dislocation measure. It should be plain from the Poissonian construction that if we denote by  $\zeta$  the instant (i.e. the first coordinate) of the first atom  $(t_{i,m}, \bar{\pi}^{(i,m)}, 1)$  with  $\bar{\pi}^{(i,m)} = (\pi^{(i,m)}, \mathbf{i})$  and  $i_1 \neq i$ , then  $\zeta$  is the first jump time of the type process  $J$  and the process killed at time  $\zeta$ ,  $(|\Pi_1(t)|, u < \zeta)$ , can be viewed as the process of the tagged fragment in a homogeneous monotype fragmentation with no erosion and dislocation measure  $\nu_i^\dagger$ . This yields

$$\Psi^{(i)}(\theta) = \Psi^{(i,\dagger)}(\theta) - \Psi^{(i,\dagger)}(0),$$

where, according to Theorem 3.2 in [6] on its page 135,

$$\Psi^{(i,\dagger)}(\theta) = \int_{\mathcal{P}_m} \left( 1 - \sum_{n=1}^{\infty} x_n^{1+\theta} \right) \nu_i^\dagger(d\mathbf{x}) = \int_{\bar{\mathcal{P}}_m} \left( 1 - \sum_{n=1}^{\infty} \mathbb{1}_{\{i_n=i\}} x_n^{1+\theta} \right) \nu_i(d\bar{\mathbf{x}}).$$

Hence

$$\Psi^{(i)}(\theta) = \int_{\bar{\mathcal{P}}_m} \left( \sum_{n=1}^{\infty} \mathbb{1}_{\{i_n=i\}} (x_n^{1+\theta} - x_n) \right) \nu_i(d\bar{\mathbf{x}}).$$

Putting the pieces together in (14), this establishes our claim.  $\square$

If we introduce

$$\underline{\theta} = \inf \left\{ \theta \in \mathbb{R} : \int_{\bar{\mathcal{P}}_m} \left| \mathbb{1}_{\{i=j\}} - \sum_{n=1}^{\infty} x_n^{1+\theta} \mathbb{1}_{\{i_n=j\}} \right| \nu_i(d\bar{\mathbf{x}}) < \infty \text{ for every } i, j \in \{1, \dots, k\} \right\},$$

then the Bernstein matrix function  $\theta \rightarrow \Phi(\theta)$  possesses an analytic extension to  $]\underline{\theta}, \infty[$  and Theorem 3 still holds for  $\theta \in ]\underline{\theta}, \infty[$ .

### 4.3 Connexion with multitype branching random walks

The preceding analysis of the evolution of the tagged fragment provides us with the key to shift some deep results on multitype branching random walks to homogeneous fragmentations. The approach is quite similar to that in [7], so again we shall skip details.

Just as in the monotype case, we consider the logarithms of the masses of the fragments and introduce for every  $t \geq 0$  the empirical measure  $\mathbf{Z}^{(t)} = (Z_1^{(t)}, \dots, Z_k^{(t)})$ , where

$$Z_j^{(t)} = \sum_{n=1}^{\infty} \mathbb{1}_{\{T_n(t)=j\}} \delta_{-\ln X_n(t)}. \quad (16)$$

For every fixed step-parameter  $a > 0$ , the process in discrete time  $(\mathbf{Z}^{(an)}, n \in \mathbb{Z}_+)$  is then a multitype branching random walk; see [9] for a precise definition. For the sake of simplicity, we shall focus on the case when  $a = 1$  and compute first a quantity of fundamental importance in terms of the characteristics of the fragmentation.

The analysis of multitype branching random walks relies on the Laplace transform of the intensity

$$m_{ij}(\theta) = \mathbb{E}_i \left( \int_{\mathbb{R}} e^{-\theta x} Z_j^{(1)}(dx) \right) = \mathbb{E}_i \left( \sum_{n=1}^{\infty} \mathbb{1}_{\{T_n(1)=j\}} X_n^\theta(1) \right).$$

Recall now that  $\bar{X}(t) = |\bar{\Pi}(t)|^\downarrow$  and, from (4), that conditionally on  $\bar{X}(t)$ , the tagged fragment  $|\bar{\Pi}_1(t)| = (\exp -S_t, J_t)$  is distributed as a size-biased sample of  $\bar{X}(t)$ . Hence, for every  $\theta > \underline{\theta} + 1$ , we have

$$m_{ij}(\theta) = \mathbb{E}_{i,0}(\exp(-(\theta - 1)S_1), J_1 = j) = (e^{-\Phi(\theta-1)})_{ij}.$$

We shall now make a further assumption on the fragmentation  $\bar{X}$ , which will be crucial to investigate its asymptotic behavior. Specifically, we assume henceforth that the process  $J$  of the type of the tagged fragment is ergodic, i.e. the intensity matrix  $\mathbf{\Lambda} = (\lambda_{ij})$  given by (15) is irreducible. We recall from the Perron-Frobenius theory (see for instance Seneta [15] or Section I.6 and II.4 in Asmussen [1]) that for every  $\theta > \underline{\theta}$ , the matrix  $\exp(-\Phi(\theta))$  has a unique real eigenvalue with maximal modulus which can be expressed as  $e^{-\varphi(\theta)}$ . In other words,  $\varphi(\theta)$  is the eigenvalue of the Bernstein matrix  $\Phi$  with minimal real part. We also write  $\mathbf{u}(\theta) = (u_1(\theta), \dots, u_k(\theta))$  and  $\mathbf{v}(\theta) = (v_1(\theta), \dots, v_k(\theta))$  for the left and right eigenvectors<sup>3</sup> associated with  $e^{-\varphi(\theta)}$ , normalized so that  $\sum_{i=1}^k u_i(\theta) = 1$  and  $\sum_{i=1}^k u_i(\theta)v_i(\theta) = 1$ .

We are now able to turn our attention to a fundamental family of martingales, which have been introduced first by Biggins in the monotype situation.

**Theorem 4** *Assume that the erosion coefficients  $c_i$  are all zero, that the dislocation measures  $\nu_i$  are conservative (in the sense that (12) holds), and that the intensity matrix (15) is irreducible.*

(i) *The equation*

$$\varphi(\theta) = (\theta + 1)\varphi'(\theta)$$

*possesses a unique solution  $\bar{\theta} \geq 0$ . The function  $\theta \rightarrow \varphi(\theta)/(\theta + 1)$  increases on  $]\underline{\theta}, \bar{\theta}[$  and decreases on  $]\bar{\theta}, \infty[$ , and thus reaches its unique maximum at  $\bar{\theta}$ .*

(ii) *For every  $\theta \in ]\underline{\theta}, \bar{\theta}[$ , the process*

$$M(\theta, t) = e^{t\varphi(\theta)} \sum_{n=1}^{\infty} v_{T_n(t)}(\theta) X_n^{\theta+1}(t), \quad t \geq 0$$

*is a martingale which converges a.s. and in  $L^1(\mathbb{P}_i)$  for every type  $i \in \{1, \dots, k\}$ . Further, this convergence is uniform for  $\theta$  in any compact set in  $]\underline{\theta}, \bar{\theta}[$ , almost surely.*

**Proof:** (i) It can be shown that the function

$$\varphi : ]\underline{\theta}, \infty[ \rightarrow \mathbb{R} \text{ is concave, increasing, and } \varphi(\theta) = o(\theta) \text{ as } \theta \rightarrow \infty. \quad (17)$$

---

<sup>3</sup>It may be useful to compare our notation with that in [9], see in particular Theorem 1 there. The matrix  $M(\theta)$ , the eigenvalue  $\rho(\theta)$  and the eigenvectors  $u(\theta)$  and  $v(\theta)$  there coincide respectively with  $\exp(-\Phi(\theta - 1))$ ,  $\exp(-\varphi(\theta - 1))$ ,  $\mathbf{u}(\theta - 1)$  and  $\mathbf{v}(\theta - 1)$  here.

See Theorem 3.7 in Seneta [15] for the concavity assertion, the fact that  $\varphi$  increases is similar. Finally observe from Theorem 3 that  $\lim_{\theta \rightarrow \infty} \theta^{-1} \Phi(\theta) = 0$ , which entails  $\varphi(\theta) = o(\theta)$ . We can then follow the arguments of the proof of Lemma 1 in [4].

(ii) We start by recall from the size-biased sampling formula (4) that, if  $(\mathcal{F}_t)_{t \geq 0}$  denotes the natural filtration of  $\bar{\mathbf{X}}$ , then

$$\begin{aligned} M(\theta, t) &= e^{t\varphi(\theta)} \sum_{n=1}^{\infty} v_{T_n(t)}(\theta) X_n^{\theta+1}(t) \\ &= e^{t\varphi(\theta)} \mathbb{E}(\exp(-\theta S_t) v_{J_t}(\theta) \mid \mathcal{F}_t) . \end{aligned}$$

As  $\mathbf{v}(\theta)$  is a right eigenvector of  $\exp(-t\Phi(\theta))$  corresponding to the eigenvalue  $e^{-t\varphi(\theta)}$ , we easily see from the Markov property of  $(J, S)$  that the process  $e^{t\varphi(\theta)} \exp(-\theta S_t) v_{J_t}(\theta)$  is a martingale in its own filtration. By projection on  $(\mathcal{F}_t)_{t \geq 0}$ , we conclude that  $M(\theta, t)$  is an  $(\mathcal{F}_t)$ -martingale.

By an argument of discretization analogous to that in [7], it suffices to establish the statement when  $t$  goes to infinity along the sequence  $an$  for some arbitrary  $a > 0$ . For the sake of simplicity, we shall focus on the case  $a = 1$  without loss of generality and aim at applying Theorems 2 and 3 of [9] to the discrete time martingale

$$M(\theta, n) = e^{n\varphi(\theta)} \sum_{j=1}^k v_j(\theta) \int e^{-(\theta+1)x} Z_j^{(n)}(dx), \quad n \in \mathbb{Z}_+.$$

Recall from Theorem 1(ii) in [9] that  $v_j(\theta) \neq 0$  for every  $\theta > \underline{\theta}$  and  $j = 1, \dots, k$ . An application of the conditional Jensen's inequality to the identity

$$M(\theta, 1) = e^{\varphi(\theta)} \mathbb{E}(\exp(-\theta S_1) v_{J_1}(\theta) \mid \mathcal{F}_1)$$

shows that for every  $\theta > \underline{\theta}$ , there is  $\alpha > 1$  such that  $\mathbb{E}_i(M(\theta, 1)^\alpha) < \infty$  for all types  $i = 1, \dots, k$ . On the other hand, we deduce from (i) that whenever  $\theta \in ]\underline{\theta}, \theta[$ , we can find  $\alpha > 1$  close to 1 such that  $\varphi(\theta)/(\theta + 1) < \varphi(\theta')/(\theta' + 1)$  where  $\theta' = \alpha(\theta + 1) - 1$ . This implies that

$$\exp(-\varphi(\alpha(\theta + 1) - 1) + \alpha\varphi(\theta)) < 1,$$

and Theorems 2 and 3 in [9] now entails our claim.  $\square$

Let us mention an interesting consequence of Theorem 4 to the rate of decay of the largest fragment as time goes to infinity. It follows readily from Theorem 4 that a.s.

$$\lim_{t \rightarrow \infty} \frac{1}{t} \ln X_1(t) = -\varphi'(\bar{\theta}),$$

see e.g. Corollary 1 in [5]. More precisely, the latter easily entails that a.s. for every type  $j \in \{1, \dots, k\}$ ,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \ln \xi_j(t) = -\varphi'(\bar{\theta}),$$

where  $\xi_j(t) = \max\{X_n(t), T_n(t) = j\}$ .

## 4.4 Asymptotic behavior of the empirical measure

We shall now conclude this section by presenting a couple of applications to the asymptotic behavior of homogeneous multitype mass-fragmentations. The first belongs to the same vein as Corollary 3.3 in [6] on its page 158. In the monotype case, a version of the result in discrete time goes back to Kolmogorov [13], in probably the first rigorous work ever on fragmentation processes. Roughly speaking, Kolmogorov provided an explanation to the fact which has been observed experimentally in mineralogy, that the logarithms of the masses of mineral grains are often normally distributed. We shall see that a similar feature holds for the more general model of multitype fragmentations.

**Corollary 1** *Assume that the erosion coefficients  $c_i$  are all zero, that the dislocation measures  $\nu_i$  are conservative (in the sense that (12) holds), and that the intensity matrix (15) is irreducible. Denote for simplicity by  $\mathbf{u}(0) = \mathbf{u} = (u_1, \dots, u_k)$  the stationary distribution on  $\{1, \dots, k\}$  of the Markov chain  $J$ , i.e.  $\mathbf{u}$  is the unique probability vector with*

$$\mathbf{u}\Lambda = \mathbf{0}.$$

*Suppose further that  $\varphi$  is twice differentiable at 0. Then the following limits hold in  $L^2(\mathbb{P}_i)$  for any initial type  $i \in \{1, \dots, k\}$  and every continuous bounded function  $f : \mathbb{R} \times \{1, \dots, k\} \rightarrow \mathbb{R}$ :*

$$\lim_{t \rightarrow \infty} \sum_{n=1}^{\infty} X_n(t) f(t^{-1} \ln X_n(t), T_n(t)) = \sum_{j=1}^k u_j f(-\varphi'(0), j)$$

and

$$\lim_{t \rightarrow \infty} \sum_{n=1}^{\infty} X_n(t) f(t^{-1/2} (\ln X_n(t) + \varphi'(0)t), T_n(t)) = \sum_{j=1}^k u_j \mathbb{E}(f(\mathcal{N}(0, -\varphi''(0)), j)),$$

where  $\mathcal{N}(0, -\varphi''(0))$  denotes a centered Gaussian variable with variance  $-\varphi''(0)$ .

Informally, the first limit means that the masses of most fragments decay exponentially fast with rate  $\varphi'(0)$  and their types are distributed according to the stationary law  $\mathbf{u}$  of the Markov chain  $J$ . The second limit is a refinement of the first and shows that, in a pathwise sense, fluctuations are Gaussian and independent of the type.

**Proof:** The two limits can be established by first and second moments estimates which rely respectively on the law of large numbers and the central limit theorem for the Markov additive process  $(S, J)$ , and an argument of propagation of chaos. For the sake of conciseness, we shall focus on the second limit.

Under the present assumptions, we know from Corollary 2.8 in [1] on its page 313 that as  $t \rightarrow \infty$ ,

$$\frac{S_t - \varphi'(0)t}{\sqrt{t}} \Rightarrow \mathcal{N}(0, -\varphi''(0)),$$

where  $\Rightarrow$  is used as a symbol for convergence in distribution. On the other hand, we also have  $J(t) \Rightarrow \tau$ , where  $\tau$  is a random type distributed according to the stationary law  $\mathbf{u}$ . Further an easy argument using the fact that the Markov chain  $J$  mixes exponentially fast shows the asymptotic independence, in the sense that

$$\left( \frac{S_t - \varphi'(0)t}{\sqrt{t}}, J_t \right) \Rightarrow (\mathcal{N}(0, -\varphi''(0)), \tau),$$

where in the right-hand side, the variables  $\mathcal{N}(0, -\varphi''(0))$  and  $\tau$  are assumed independent.

Recall now that  $\bar{X}(t) = |\bar{\Pi}(t)|^\downarrow$  and, from (4), that conditionally on  $\bar{X}(t)$ , the tagged fragment  $|\bar{\Pi}_1(t)| = (\exp -S_t, J_t)$  is distributed as a size-biased sample of  $\bar{X}(t)$ . Hence

$$\mathbb{E}_i \left( \sum_{n=1}^{\infty} X_n(t) f(t^{-1/2}(\ln X_n(t) + \varphi'(0)t), T_n(t)) \right) = \mathbb{E}_i \left( f \left( \frac{-S_t + \varphi'(0)t}{\sqrt{t}}, J_t \right) \right),$$

and therefore

$$\begin{aligned} & \lim_{t \rightarrow \infty} \mathbb{E}_i \left( \sum_{n=1}^{\infty} X_n(t) f(t^{-1/2}(\ln X_n(t) + \varphi'(0)t), T_n(t)) \right) \\ &= \mathbb{E} (f(\mathcal{N}(0, -\varphi''(0)), \tau)) \\ &= \sum_{j=1}^k u_j \mathbb{E}(f(\mathcal{N}(0, -\varphi''(0)), j)). \end{aligned}$$

By an argument of propagation of chaos similar to that in the proof of Corollary 3.3 in [6] on its page 159-160, we can estimate the second moment and get

$$\begin{aligned} & \lim_{t \rightarrow \infty} \mathbb{E}_i \left( \left( \sum_{n=1}^{\infty} X_n(t) f(t^{-1/2}(\ln X_n(t) + \varphi'(0)t), T_n(t)) \right)^2 \right) \\ &= \left( \sum_{j=1}^k u_j \mathbb{E}(f(\mathcal{N}(0, -\varphi''(0)), j)) \right)^2. \end{aligned}$$

This entails the convergence in  $L^2(\mathbb{P}_i)$  which has been stated.  $\square$

Finally, using time discretization techniques similar to those in [7], we can translate Theorem 7 in [9] to multitype fragmentations. This yields a pathwise large deviation limit theorem for the empirical distribution of the fragmentation which refines considerably Corollary 1. In this direction, we shall further assume that the eigenvalue function  $\varphi$  is strictly concave and that the branching random walk  $\mathbf{Z}^{(n)}$  is strongly non-lattice (see [9] for the terminology), which are very natural and mild conditions. Recall Theorem 4 and denote the terminal value of the martingale  $M(\theta, t)$  by  $M(\theta, \infty)$ .

**Corollary 2** *Let  $h : \mathbb{R} \times \{1, \dots, k\} \rightarrow \mathbb{R}$  be a continuous function with compact support. Under the preceding assumptions, we have*

$$\begin{aligned} & \lim_{t \rightarrow \infty} \sqrt{t} e^{-t((\theta+1)\varphi'(\theta) - \varphi(\theta))} \sum_{n=1}^{\infty} h(t\varphi'(\theta) + \ln X_n(t), T_n(t)) \\ &= \frac{M(\theta, \infty)}{\sqrt{2\pi|\varphi''(\theta)|}} \sum_{j=1}^k u_j(\theta) \int_{-\infty}^{\infty} h(y, j) e^{-(\theta+1)y} dy, \end{aligned}$$

*uniformly for  $\theta$  in compact subsets of  $]\underline{\theta}, \bar{\theta}[$ , almost surely.*

In particular, this implies that for every  $a < b \in \mathbb{R}$ ,  $\theta \in ]\underline{\theta}, \bar{\theta}[$  and  $j \in \{1, \dots, k\}$ , there is the estimate as  $t \rightarrow \infty$

$$\begin{aligned} & \# \left\{ n \in \mathbb{N} : ae^{-t\varphi'(\theta)} \leq X_n(t) \leq be^{-t\varphi'(\theta)} \text{ and } T_n(t) = j \right\} \\ & \sim A_\theta u_j(\theta) t^{-1/2} \exp(t((\theta+1)\varphi'(\theta) - \varphi(\theta))) (e^{-a(\theta+1)} - e^{-b(\theta+1)}), \end{aligned}$$

where  $A_\theta$  is some strictly positive random variable with finite mean. See Corollary 2 in [9].

**Agradecimentos :** In August 2004, I had the pleasure to be invited in Ubatuba, to give a short course on self-similar fragmentation chains for the 8th Brazilian School of Probability. The present paper can be viewed, in some sense, as a natural prolongation of the material which I presented during that course. I would like to thank again Pablo Ferrari for his very kind invitation and the wonderful organization, and Vladas Sidoravicius and Maria-Eulalia Vares for making my stays so pleasant each time I have the chance to go to Rio.

## References

- [1] Asmussen, S. (2003). *Applied Probability and Queues*. Second edition. Applications of Mathematics. Stochastic Modelling and Applied Probability. Springer-Verlag, New York.
- [2] Barral, J. (2001). Generalized vector multiplicative cascades. *Adv. in Appl. Probab.* **33**, 874-895.
- [3] Berestycki, J. (2002). Ranked fragmentations. *ESAIM, Probabilités et Statistique* **6**, 157-176. Available via <http://www.edpsciences.org/ps/OnlinePSbis.html>
- [4] Bertoin, J. (2003). The asymptotic behavior of fragmentation processes. *J. Euro. Math. Soc.* **5**, 395-416.
- [5] Bertoin, J. (2006). Some aspects of a random fragmentation model. *Stochastic Process. Appl.* **116**, 345-369.

- [6] Bertoin, J. (2006). *Random Fragmentation and Coagulation Processes*. Cambridge University Press, Cambridge.
- [7] Bertoin, J. and Rouault, A. (2005). Discretization methods for homogeneous fragmentations. *J. London Math. Soc.* **72**, 91-109.
- [8] Biggins, J.D. and Kyprianou, A.E. (2004). Measure change in multitype branching. *Adv. Appl. Probab.* **36**, 544-581.
- [9] Biggins, J.D. and Rahimzadeh Sani, A. (2005). Convergence results on multitype, multivariate branching random walks. *Adv. Appl. Probab.* **37**, 681-705.
- [10] Dong, R., Gnedin, A., and Pitman, J. (2006). Exchangeable partitions derived from Markovian coalescents. Preprint available via : <http://arxiv.org/abs/math.PR/0603745>
- [11] Haas, B., Miermont, G., Pitman, J., and Winkel, M. (2006). Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. Preprint available via <http://arxiv.org/abs/math.PR/0604350>
- [12] Kingman, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13**, 235-248.
- [13] Kolmogoroff, A. N. (1941). Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *C. R. (Doklady) Acad. Sci. URSS* **31**, 99-101.
- [14] Pitman, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27**, 1870-1902.
- [15] Seneta, E. (1973). *Non-Negative Matrices. An Introduction to Theory and Applications*. Halsted Press, New York.